

分类号: TM863

单位代码: 10335

学 号: 11621033

# 浙江大学

## 硕士学位论文



中文论文题目: 基于电网仿真数据的  
异常检测与可视分析

英文论文题目: **English title**

申请人姓名: 李宗壮

指导教师: 各类 网络文档

专业名称: 计算机科学与技术


研究方向: 可视化

所在学院: 计算机科学与技术学院

论文提交日期 2015 年 4 月 8 日

## 浙江大学研究生学位论文独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得 浙江大学 或其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。


学位论文作者签名： 


签字日期： 2015 年 6 月 30 日

## 学位论文版权使用授权书

本学位论文作者完全了解 浙江大学 有权保留并向国家有关部门或机构送交本论文的复印件和磁盘，允许论文被查阅和借阅。本人授权 浙江大学 可以将学位论文的全部或部分内容编入有关数据库进行检索和传播，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

（保密的学位论文在解密后适用本授权书）

学位论文作者签名： 

导师签名： 

签字日期： 2015 年 6 月 30 日

签字日期： 2015 年 6 月 30 日

## 摘 要

由于电网系统的不稳定造成的严重停电问题可导致巨大的经济损失，所以研究电力系统的稳定性一直是电力行业的一个核心问题。其中电网数据的异常检测和电网的故障检测占据着重要的地位。传统上，电力行业工作者仅是对电力系统的稳定性进行定量的分析，难以以更为高效的方式来进行。因此，本文作者所在团队与中国南方电网合作，着力开发构建了对电网数据进行异常检测和故障检测的可视系统。本文所描述的工作具有三个重要贡献：（1）设计并实现了一款支持用户对电网的异常及故障进行识别的可视分析系统；（2）将视觉编码与统计质量控制方法相结合，提供了一种能够考察多个变量对系统稳定性所造成的影响总和的研究方法；（3）一套用于多变量时变数据、能使分析人员随时检测电网异常，具体分析故障的可视系统本文主要电网数据的异常与故障检测和可视系统的实现和测试两方面介绍本文作者在本文系统开发过程中所做的贡献。

**关键词：** 异常检测，故障检测，可视分析，电力网络，统计质量控制，深度神经网络

## Abstract

Because the serious blackouts caused by the instability of power grid system can lead to huge economic losses, the study of power system stability has always been a core issue in the power industry. Power grid data anomaly detection and power grid fault detection occupy an important position.

Traditionally, the power industry workers only make quantitative analysis of the stability of the power system, which is difficult to carry out in a more efficient way. Therefore, the author's team cooperated with China Southern Power Grid to develop and construct a visual system for anomaly detection and fault detection of power grid data. The work described in this paper has three important contributions: (1) Design and implement a visual analysis system to support users to identify anomalies and faults in power grids; (2) Combine visual coding with statistical quality control method to provide a total of effects of multiple variables on system stability. Research methods; (3) A visual system for multi-variable time-varying data, which can enable analysts to detect abnormal power grids at any time and analyze specific faults.

This paper mainly introduces the contribution of the author in the development of the system from the aspects of anomaly and fault detection of power grid data and the realization and testing of the visual system.

**Keywords:** Anomaly detection, fault detection, visual analysis, power network, statistical quality control, deep neural network

插 图

1-1 标准电网格式 ..... 1

表 格

# 目 次

摘要 .....	I
Abstract.....	II
插图 .....	III
表格 .....	IV
目次 .....	
1 绪论 .....	1
1.1 课题背景 .....	1
1.1.1 什么是电网.....	1
1.1.2 异常的定义，电网异常是什么，本文中检测的异常指什么.....	2
1.2 可视分析的作用，引入可视分析的原因 .....	2
1.3 相关工作 .....	3
1.3.1 异常检测及异常检测可视分析.....	3
1.3.2 电网异常检测.....	3
1.3.3 电网故障检测.....	3
1.3.4 时序多变量数据可视分析.....	3
1.4 本文工作 .....	3
1.5 章节安排 .....	3
2 数据与分析任务 .....	5
2.1 数据 .....	5
2.1.1 仿真数据.....	5
2.1.2 电量数据.....	6
2.2 分析任务 .....	6
2.2.1 异常检测任务.....	6
2.2.2 故障检测任务.....	6

3	基于统计质量控制方法的异常检测 .....	7
3.1	控制图与统计质量控制 .....	7
3.1.1	控制图理论的基本思想 .....	7
3.1.2	统计质量控制的基本思想 .....	8
3.2	已有统计质量控制图分析 .....	8
3.2.1	介绍已有的统计质量控制图方法各自的特点特征 .....	8
3.2.2	结合电网数据特色与分析任务, 综合说明选择 shewhart 控制图的原因 .....	8
3.3	异常检测算法 .....	8
3.3.1	shewhart 统计质量控制方法原理 .....	8
3.3.2	基于 shewhart 方法改进的异常检测算法 .....	8
3.4	实验结果 .....	8
3.4.1	算法复杂度、效率 .....	8
3.4.2	检测结果分析 .....	8
4	故障检测 .....	9
4.1	故障检测问题 .....	9
4.1.1	具体问题说明 .....	9
4.1.2	问题抽象 (如何抽象为一个分类问题) .....	9
4.2	深度神经网络模型 .....	9
4.2.1	选择深度神经网络的原因 .....	9
4.2.2	RNN 原理 .....	9
4.2.3	为何选择基于 RNN 的 LSTM 模型 .....	9
4.2.4	本文所用模型 (LSTN-FCN) 的网络结构与原理 .....	9
4.3	具体实现 .....	9
4.3.1	数据处理 .....	9
4.3.2	参数设置 (调参过程) .....	9
4.3.3	实现 (keras 库、基于 tensorflow、变成语言 python) .....	9
5	系统设计与可视设计 .....	10
5.1	系统实现 .....	10
5.1.1	系统结构 .....	10
5.1.2	后端设计 .....	10
5.1.3	前端设计 .....	10
5.2	可视设计 .....	10
5.3	实验结果与案例分析 .....	10



6	总结与展望 .....	11
6.1	讨论 .....	11
6.2	未来工作 .....	11
6.3	总结 .....	11

# 1 绪论

## 1.1 课题背景

### 1.1.1 什么是电网

电网是电力系统的典型形式，描述为通过组织良好的配电系统从发电机向用户传输电力的网络。电力系统是一个实时运行的物理系统，其中含有大量发电机、线路、变压器，以及负荷等不同类型的、不同特性的元件。为了保障电力系统的安全运行、避免停电事故，需要对电网的运行作出预先安排，并对电网发生的故障和重要事件进行及时分析。电网的典型表示是用节点代表母线，而节点之间的连线则代表连接两条母线的电力线，特殊的，带有闪电标志的节点代表了发电机，如图 1-1 所示。电网是复杂的、动态的、多方面的系统。

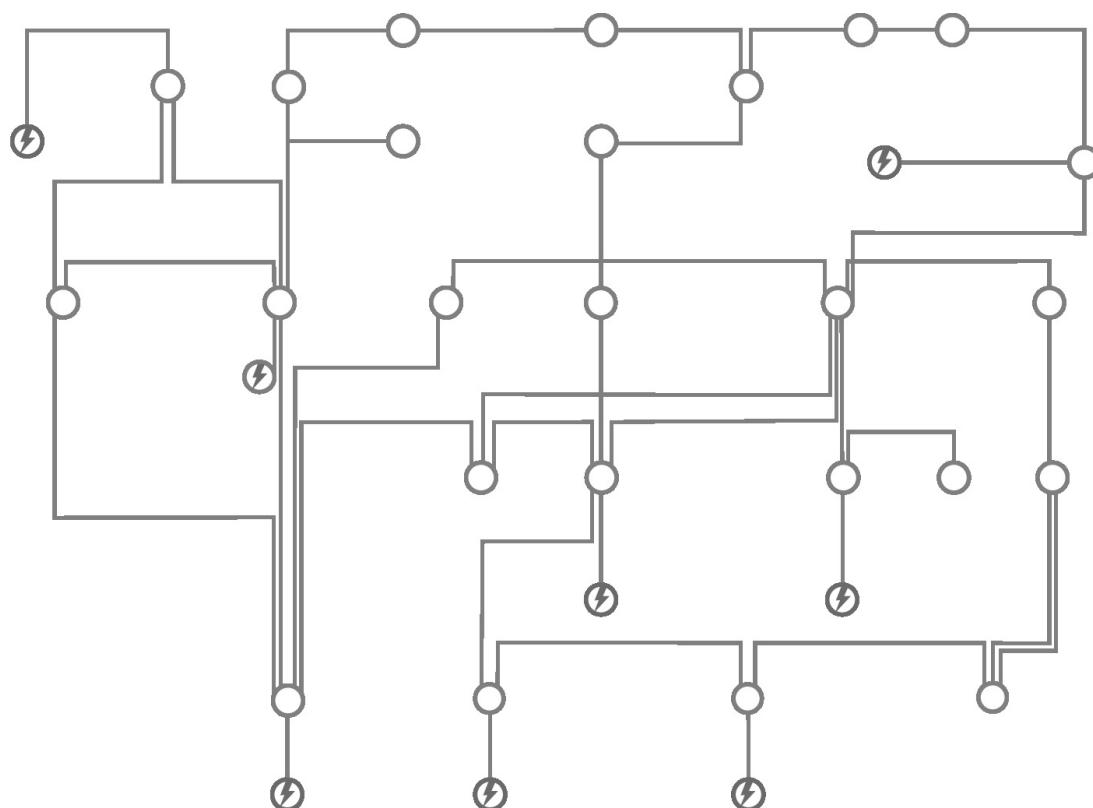


图 1-1 标准电网格式

### 1.1.2 异常的定义，电网异常是什么，本文中检测的异常指什么

本文基于数值模拟数据，特别是暂态稳定数据，对电网稳定性进行了可视化分析。特别地，将统计质量控制（SQC）方法集成到具有创新视觉设计的稳定性时间的检测和研究中。对基于 SQC 的方法的增强促进了细粒度的稳定性判断，而细粒度的稳定性判断又通过富有表现力的视觉证据得到增强。将 SQC 方法、可视化编码、领域知识和领域专家的经验无缝结合视为推进智能电网的新机会。我们的工作密切合作领域专家和经验丰富的电力仿真运营商。它们提供领域经验、程序、数据集和分析任务。最终设计和实现了一个综合的视觉系统，以支持多时间序列的视觉探索和分析，识别和区分显著的不稳定模式。支持电网的异常、不稳定和故障的勘探、识别和推理的可视化分析系统。视觉编码和 SQC 方法的系统集成，为研究可能影响系统稳定性的多个相关因素的联合效应提供可靠的视觉证据。一套用于数字时间序列的视觉设计，使分析人员具有对全球趋势、局部变化和多变量随时间变化的相关性的视觉检查的能力。

## 1.2 可视分析的作用，引入可视分析的原因

在人类的各类感觉器官中，人眼具有最强的信号处理能力。因此视觉是获取信息的最重要通道，超过 50% 可视化是人类在认知过程中产生的对事物的认知视图，易于理解并有利于知识交流。而信息可视化是让计算机绘制动态、交互的外部辅助视图，目的是增强人的认知能力。

随着数据规模的增加，直接呈现数据会导致显示空间不足、图元遮挡、绘制效率低等问题。另外，现代分析任务更加复杂，仅仅依赖可视化难以令人发现隐藏在数据背后的模式，迭代式交互分析是领域专家认识和理解数据的重要方法。为此，可视化专家提出了可视分析概念，它是一种通过交互式可视化界面来辅助用户对大规模复杂数据集进行分析推理的科学和技术。可视分析不仅仅是可视化，它集成了可视化、人的要素（如交互、认知、感知、合作、表示和传播）和数据分析等。相对于可视化，可视化分析整合了信息分析、地理空间分析和科学分析等技术，而且人的要素在人与机器的通信过程和决策过程中起到非常重要的作用。可视分析是人的智能（知识和经验）和机器智能（计算）多次协作的迭代过程。

可视分析有效融合了计算机强大的计算能力和人的认知能力，能够发挥二者的优势完成各自擅长的任务。因此，可视分析是交叉融合了可视化、人机交互、认知科学、数据挖掘、信息论、决策理论等研究领域的新研究方向。

可视分析需要合适的数据载体，而电力系统数据便是合适的载体。电力系统稳定是电力系统中的一个基本问题。电力工业因系统不稳定性会引起严重的停电事故和巨大的经济损失。电网是电力系统的一种经典形式，被描述为通过发电机向用户传输电力的网络，一个结构良好的分配制度。一个电网是复杂的、动态的、多方面的系统。直到现在还没有一种可行的方法来纪录系统的情况。控制实验对实际的电力系统可能会造成意外的损失。这使得模拟计算性能分析成为唯一可以接受的方法。目前在实际应用中已经验证，模拟计算可以达到比较精确的测试。定期运行数据模拟，对实际事件进行推理，预防性问题分析。

在可视分析中所提供的数据驱动的分析工具，能够有助于分析人员从更多地角度、更深入地探索海量仿真数据中所蕴含的规律，使得电网运行分析具有更强地研究性，适应我国电网高速发展的现状。

## 1.3 相关工作

### 1.3.1 异常检测及异常检测可视分析

异常检测在分析动态系统中占据着及其重要的地位，特别地，可视化为识别和发现异常提供了视觉帮助。不同的领域需要各种可视化手段，以适应数据特征和领域问题的具体特征。

### 1.3.2 电网异常检测

### 1.3.3 电网故障检测

### 1.3.4 时序多变量数据可视分析

大数据时代，数据的复杂性与日俱增。为了更好的表现出时序多变量随着时间序列的数量和复杂性的增加，在视觉分析中需要比较手段。一般的，信息可视化的比较设计分为以下几类。

并列法（Juxtaposition）按照时间（空间）的顺序对时间序列数据进行排列。比较典型的有如图??所示的 LeadLine 和如图??CloudLines，它们通过在原子级显示多个时间序列而不进行聚合，从而将并置方法扩展到多个时间序列的大规模数据。

交叠法（Superposition）在同一空间中覆盖多个对象，通过比较具有不同颜色或透明度的层，用户可以直接观察对象之间的差异。Collberg 等人使用了并置法将多张研究对象的快照放置在同一空间中以显示图随时间发生的变化，如图??所示。

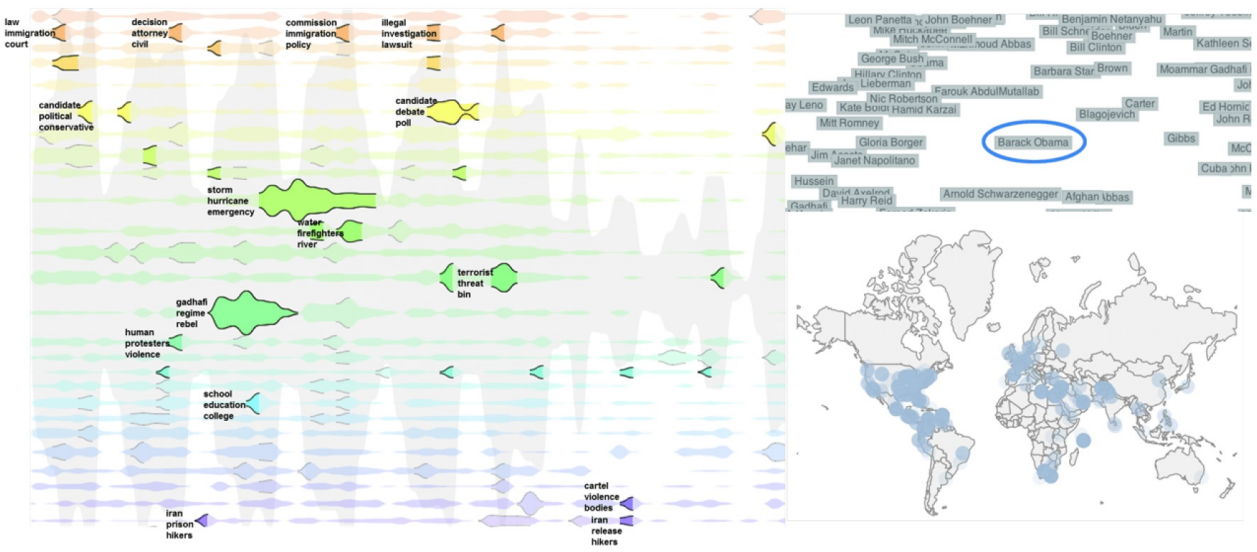


图 1-2 Leadlines

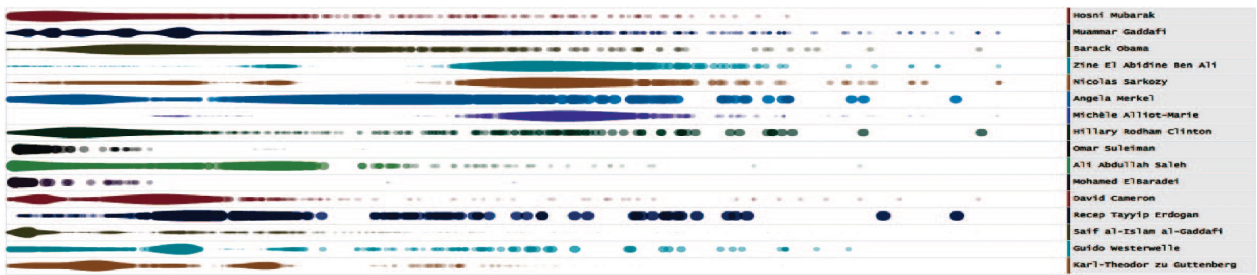


图 1-3 Cloudlines

显式编码法（Explicit Encoding）利用时间序列数据的不同视觉编码来揭示对象之间的关系或对比。World Lines 是多重集成仿真比较可视化的一个复杂例子，其视觉设计结合了包括了上述的三种设计，表现如图??所示。

1.4 本文工作

本项目根据获得的相关电网数据，包括电网各个关键节点（母线）的电压值、频率值等，以及发电机之间的相对功角值，对电网的异常和故障采取合适的方式进行检测。同时还会将这些内容以可视化的方式简单的呈现出来，来方便相关研究人员的理解和处理。最后，我们会使用数据来验证用到的所有算法的合理性与准确性。

1.5 章节安排

本文第二章介绍了电网数据并进行了关于异常故障检测的任务分析，以便于进行下一步的研究；第三章便是根据任务选取了统计质量控制的方法来进行异常检测，并对现有的

统计质量控制方法进行了具体的介绍；第四章是对于故障检测的处理，选用了 RNN 的深度神经网络模型，并给出了具体的实现和相应的结果；第五章便是整体系统的呈现，包括前端设计，后端设计等诸多方面，还对具体的案例进行了分析；第六章包括了作者对整个系统的总结和未来的展望。

## 2 数据与分析任务

### 2.1 数据

#### 2.1.1 仿真数据

仿真数据，尤其是暂态稳定计算在指导电网稳定运行中起着重要的作用。仿真数据通常是描述每个节点中每个变量的演变的一组时间序列。在模拟它的运行、故障和恢复过程中，不同形式的不稳定性伴随着它的演化过程而出现。数值模拟方法如李亚普诺夫函数及其变分已被证明是研究动态和瞬态不稳定性的独特手段。因而，分析人员要识别故障事件的发生，研究故障事件的类型和模式，并通过分析变量的时间序列来推断故障事件的原因，是非常麻烦的。目前缺乏定量描述电力系统稳定状态的理论模型。传统的统计方法，如统计质量控制，提供了基于点的稳定性判断，但对于刻画全局趋势和模糊变化是低效的。

电力系统仿真分析广泛应用于电力系统运行、规划、设计等领域，是电力系统的一项基础性支撑技术。经过数十年的发展，我国基于机理模型的电力系统仿真计算技术已处于世界前列，已有 PSD、PSASP、ADPSS、PS-Model，以及全过程仿真软件等诸多广泛应用的自主知识产权优秀产品，可以基本满足大电网仿真分析的需求。

电力系统仿真分析中的“仿真”是指对电网元件进行数学建模，使用微分-代数方程组表达各个元件的外特性，通过求解这个方程组得到系统的状态变量值，从而掌握整个系统的情况。“分析”则是基于得到的仿真结果进行系统特性研究、控制措施制定等工作。在实际工作中，电网仿真分析可以看作是“提出假设”与“计算验证”的循环过程。传统的仿真数据分析方式有着各种限制。首先是需要耗费大量人力，电网的日益复杂，使得人力的需求大大增加；其次是十分依赖专家的经验，分析的结论与专家的经验水平息息相关，能否发现问题和能否找出解决方法都取决于专家；再就是容易出现漏洞，由于主要依赖人工分析，难免发生错漏的现象；最后是无法控制误差。

先进的仿真计算技术必须要有先进的数据分析技术和人机交互方式才能充分发挥其作用。虽然目前的各类电网仿真软件均由结果可视化和报表功能，但其着重点仅限于数据的展示，对于数据规律的分析主要还是依靠人工进行只能观察很少的重要元件和断面。这极

大地限制了对仿真结果和仿真能力的应用，使得许多在计算结果中蕴含的电网运行规律难以被发现。

### **2.1.2 电量数据**

## **2.2 分析任务**

本文的主要目标便是做出一个可以用于电网异常检测和故障检测的可视系统，用以方便相关研究人员的分析和处理。

### **2.2.1 异常检测任务**

### **2.2.2 故障检测任务**



### 3 基于统计质量控制方法的异常检测

统计质量控制（SQC）方法广泛用于监控过程，以确保合格确认产品的比例。在各种 SQC 方法中，控制图在测量过程稳定性和检测显著变化方面表现出突出的性能。在本文中，SQC 和控制图的理论构成了异常检测的基础。

#### 3.1 控制图与统计质量控制

##### 3.1.1 控制图理论的基本思想

控制图亦称“质量管理图”、“质量评估图”。根据数理统计原理分析和判断工序是否处于稳定状态所使用的、带有控制界限的一种质量管理图表。它是一种有控制界限的图，用来区分引起的原因是偶然的还是系统的，可以提供系统原因存在的资讯，从而判断生产过於受控状态。

### 3.1.2 统计质量控制的基本思想

## 3.2 已有统计质量控制图分析

### 3.2.1 介绍已有的统计质量控制图方法各自的特点特征

### 3.2.2 结合电网数据特色与分析任务，综合说明选择 **shewhart** 控制图的原因

## 3.3 异常检测算法

### 3.3.1 **shewhart** 统计质量控制方法原理

### 3.3.2 基于 **shewhart** 方法改进的异常检测算法

## 3.4 实验结果

### 3.4.1 算法复杂度、效率

### 3.4.2 检测结果分析

## 4 故障检测

### 4.1 故障检测问题

#### 4.1.1 具体问题说明

#### 4.1.2 问题抽象（如何抽象为一个分类问题）

### 4.2 深度神经网络模型

#### 4.2.1 选择深度神经网络的原因

#### 4.2.2 RNN 原理

#### 4.2.3 为何选择基于 RNN 的 LSTM 模型

#### 4.2.4 本文所用模型（LSTN-FCN）的网络结构与原理

### 4.3 具体实现

#### 4.3.1 数据处理

#### 4.3.2 参数设置（调参过程）

#### 4.3.3 实现（keras 库、基于 tensorflow、变成语言 python

## **5 系统设计与可视设计**

### **5.1 系统实现**

#### **5.1.1 系统结构**

#### **5.1.2 后端设计**

#### **5.1.3 前端设计**

### **5.2 可视设计**

### **5.3 实验结果与案例分析**

## **6 总结与展望**

### **6.1 讨论**

### **6.2 未来工作**

### **6.3 总结**